



# Working Papers

[www.cesifo.org/wp](http://www.cesifo.org/wp)

## An Explanation of (First Round) Contributions in Public-Good Experiments

Joachim Weimann  
Jeannette Brosig-Koch  
Timo Heinrich  
Heike Hennig-Schmidt  
Claudia Keser / Christian Stahr

CESIFO WORKING PAPER NO. 5039  
CATEGORY 13: BEHAVIOURAL ECONOMICS  
OCTOBER 2014

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# An Explanation of (First Round) Contributions in Public-Good Experiments

## Abstract

The dynamics of behavior observed in standard public-good experiments can be explained by imperfect conditional cooperation combined with social learning (Fischbacher and Gächter, 2010). But it is unclear what determines first-round contributions. We argue that first-round contributions depend on the difference between the actual marginal per capita return (MPCR) and the minimal MPCR necessary to create a social dilemma for the given group size. We use a novel connected-lab design that allows members of large groups to simultaneously interact under laboratory conditions. In total, 5,220 subjects participated in our experiment in groups of up to 100 subjects. The results confirm that first-round contributions increase with the MPCR distance. This effect carries over to contributions made in subsequent rounds. Our results demonstrate that small groups behave similar to large groups and that the salience of the di-lemma situation is of great importance for contributions made to the public good.

JEL-Code: C910, C720.

Keywords: public-good experiments, large group size, connected-lab design, salience of the dilemma situation, social learning.

*Joachim Weimann\**  
*University of Magdeburg*  
*Postbox 4120*  
*Germany – 39016 Magdeburg*  
*weimann@ww.uni-magdeburg.de*

*Timo Heinrich*  
*University of Duisburg-Essen*  
*Geibelstrasse 41*  
*Germany – 47057 Duisburg*  
*timo.heinrich@uni-due.de*

*Claudia Keser*  
*University of Göttingen*  
*Platz der Göttinger Sieben 3*  
*Germany – 37073 Göttingen*  
*claudia.keser@uni-goettingen.de*

*Jeannette Brosig-Koch*  
*University of Duisburg-Essen*  
*Universitaetsstrasse 12*  
*Germany – 45117 Essen*  
*jeannette.brosig@uni-due.de*

*Heike Hennig-Schmidt*  
*University of Bonn*  
*Adenauerallee 24-42*  
*Germany – 53113 Bonn*  
*hschmidt@uni-bonn.de*

*Christian Stahr*  
*University of Magdeburg*  
*Postbox 4120*  
*Germany – 39016 Magdeburg*  
*christian.stahr@ovgu.de*

October 13, 2014

We thank Ren Ohinata for research assistance. We thank Christoph Helbach, Andreas Markstädter, Martin Schmidt, and Jan Siebert for their support in running the experiments. We are grateful for the helpful comments we received at the Annual Meeting of the German Economic Association (VfS) 2013 in Düsseldorf, the Annual Conference of the German Association of Experimental Economic (GfeW) 2013 in Helmstedt, the Pacific ESA conference in Auckland 2014, and the Karlsruhe Workshop on Economics and Finance. We thank Bernd Irlenbusch, Axel Ockenfels, Bettina Rockenbach, Karim Sadrieh, and Reinhard Selten for very helpful suggestions. Financial support by Deutsche Forschungsgemeinschaft (German Research Foundation, WE 2027/10-1/2) is gratefully acknowledged.

## 1. Introduction

Experimental research in economics has demonstrated that people are willing to contribute voluntarily to the provision of public goods, even if the production function is linear and choosing a contribution of zero is the dominant strategy. As a result, subjects are partly successful in overcoming the social dilemma they are confronted with in a standard public-good experiment (SPGE).<sup>1</sup> However, this result is often accompanied by the observation that contributions in repeated public-good games decay and cooperation breaks down in the last rounds. Cooperation appears to be possible but is fragile and unstable.<sup>2</sup>

One prominent explanation for observed behavior is based on the assumption that people have social (or other-regarding) preferences (see, e.g., Thomas R. Palfrey and Jeffrey E. Prisbrey, 1997). During the last decade, experimental economists have invested a considerable amount of effort in attempting to find an answer as to why cooperation *decreases* in the course of a repeated public-good game. They have identified a particular kind of other-regarding preferences that may cause the decrease of cooperation. People like to cooperate as long as others cooperate, i.e., they behave as *conditional cooperators*.<sup>3</sup> In a recent study, Urs Fischbacher and Simon Gächter (2010) convincingly demonstrate that it is *social learning* in combination with *imperfect* conditional cooperation that drives contributions down. People learn that others contribute (and how much they contribute) and this motivates conditional cooperators to behave cooperatively as well. However, they do not necessarily match the average contributions perfectly; rather they tend to choose a contribution level slightly below the average. This imperfect matching of the average contribution is a convincing explanation for the decay of cooperation that has been observed so often in SPGEs.

The explanation given by Fischbacher and Gächter (2010) helps us to understand the dynamics of behavior in SPGEs, but an important piece is missing. Social learning about the behavior of others cannot explain the heterogeneity of first-round contribution levels observed across previous SPGEs (or, across one-shot public-good games). These first-round contribution levels are of great importance, since they are the starting point for social learning in all subsequent rounds

---

<sup>1</sup> We define standard public-good experiments as those which (a) use the voluntary contribution mechanism with a linear production function introduced by Mark Isaac, James Walker and Susan Thomas (1984) and (b) are conducted anonymously, which means that subjects have no direct contact and cannot communicate with each other.

<sup>2</sup> For an early overview see John Ledyard (1995) and for a more selective recent survey see Annanish Chaudhuri (2011). For the decay of contributions see, e.g., James Andreoni (1988, 1995), Joachim Weimann (1994), Roberto Burlando and John Hey (1997), Claudia Keser and Frans van Winden (2000), Simon Gächter and Ernst Fehr (1999), Axel Ockenfels and Joachim Weimann (1999), and Urs Fischbacher and Simon Gächter (2010).

<sup>3</sup> Claudia Keser and Frans van Winden (2000), Urs Fischbacher, Simon Gächter and Ernst Fehr (2001), Rachel Croson (2007).

and, thus, determine to a large extent the overall level of contributions (e.g., Claudia Keser and Frans van Winden, 2000, Martin Beckenkamp et al., 2013). If the average first-round contribution of group A is higher than that of group B, then it is very likely that overall contributions will be higher in A than in B. But what determines first-round contributions?

It is a natural assumption that the perception people form of the game is created by its parameters, i.e., by the payoff function given to subjects. The payoff function (including information on how many players are in the group) is all that subjects know about their decision situation when they start the experiment. Our conjecture is that the *salience* of the dilemma situation induced by this payoff function affects first-round contributions because in a social dilemma everybody profits from cooperating. This might convince conditional cooperators to make a voluntary contribution to the public good. But, as suggested by Fischbacher and Gächter (2010), conditional cooperators do not like to be the only ones to cooperate. They want to be confident that they are not exploited – and this already holds for the first round. Therefore, it is not only important that these subjects recognize that they are faced with a social dilemma (in which cooperation pays off for the group), but also that they are convinced that their group members realize, too, that they are in a dilemma situation. The more salient the dilemma situation is, the more subjects can be sure that all group members have understood the social dilemma they are in, and the more likely they are to cooperate.

We argue that the salience of a dilemma situation in a SPGE is crucially influenced by what we call “MPCR distance”. We define MPCR distance as the difference between the actual marginal per capita return (MPCR) used in the experiment and the minimal MPCR necessary to create a social dilemma. It appears to be an appropriate measure for the salience of the public good problem for the following reason. If the actual MPCR is high enough (compared to the minimal MPCR) it is not only clear that it pays off for the group to cooperate, but also that all group members understand this. That is, a high MPCR distance implies a high salience of the dilemma situation and, consequently, induces high first-round contributions. Moreover, there might exist some high level of the MPCR distance, at which all group members recognize that they are faced with a dilemma situation and at which this is common knowledge. In such a situation, the dilemma situation is fully salient and a further increase in the MPCR distance should not affect first-round contributions.

The minimal MPCR depends on the size of the group. If the group is large enough, even a small MPCR can create a high MPCR distance and, thus, a high *salience* of the dilemma situation. This implies that in large groups small changes in small MPCRs should have strong effects on

first-round contributions (as long as the dilemma situation is not fully salient yet, i.e. the MPCR distance is not too high). To test this implication, we conduct a series of experiments with large groups of 60 and 100 subjects, simultaneously playing the repeated public-good game under controlled laboratory conditions similar to those used for small groups. For both group sizes we employ a very small MPCR of 0.02, and increase this MPCR to 0.04. That is, we vary the MPCR distance for rather low values of this distance (0.003 to 0.030; see Table 1). For control purposes, we also employ groups with 30 and 40 subjects (and confront them with MPCRs of 0.04, 0.06, and 0.12, resulting in MPCR distances ranging from 0.007 to 0.095) and rather small groups with 8 subjects (and confront them with an MPCR of 0.25, resulting in an MPCR distance of 0.125). There is an additional point that makes investigating large groups confronted with very small MPCRs attractive: with this parameter constellation we increase the external validity of SPGEs, since public good problems outside the laboratory are very often characterized by the combination of large groups and very small MPCRs. We contribute to the findings of Mark Isaac, James Walker, and Arlington Williams (1994), who were the first to investigate public good provision by large groups in an experiment (but used a different design and parameter constellation; see section 2).

Our results strongly support the conjecture that the MPCR distance determines first-round contributions. In particular, a higher MPCR distance is associated with a higher average level of first-round contributions as long as the MPCR distance is not too high. For high values, we find no effect on first-round contributions if we further increase the MPCR distance. In line with the explanation given by Fischbacher and Gächter (2010), we observe that in large groups contributions also significantly decay over the course of the experiment. We provide further support for this explanation insofar as we observe less decay in contributions when giving no feedback about other group members' behavior, i.e., when preventing them from learning about others. Furthermore, this paper makes a methodological contribution to the literature: We employ a novel connected-lab design that allows the basic characteristics of conventional (one-) laboratory experiments to be maintained also with large groups.

In section 2 we introduce our research questions and conjectures referring to the literature on public-good experiments. In section 3 we present the experimental design. Section 4 includes our results and statistical analyses. Section 5 summarizes and discusses our observations.

## **2. MPCR distance**

SPGEs share a lot of common design elements. For example, they are conducted anonymously without any communication between subjects. The payoff function often varies across experiments, however. In a SPGE, the payoff function depends crucially on two parameters: the number of subjects in the group,  $N$ , and the marginal per capita return of an investment in the public good (MPCR). The public-good game is typically played repeatedly over  $T$  rounds. The total number of rounds is common knowledge and subjects are informed about the outcome after each round. Let  $z_{it}$  denote the endowment of subject  $i$  in round  $t$ ,  $b_{it}$  the individual contribution to the public good in round  $t$ , and  $p$  the marginal return on the part of the endowment not invested in the public good. The parameter  $a$  measures the marginal productivity of the public good. The payoff function for subject  $i$  in round  $t$  is then given by:

$$\Pi_{it} = (z_{it} - b_{it})p + \frac{a}{N} \sum_{j=1}^N b_{jt} \quad (1)$$

$$\text{with } a > p > \frac{a}{N} \quad (1^*) \quad \text{and } \text{MPCR} = \frac{a}{Np} \quad (1^{**})$$

Condition (1\*) ensures that (1) describes a social dilemma, which is characterized by the fact that individually rational behavior leads to an inefficient provision of the public good. If we normalize  $p=1$ , it follows that  $\text{MPCR}=a/N$  and has to be larger than  $1/N$ . Since  $1/N$  is the lower bound for the MPCR (i.e., the minimal MPCR necessary to create a social dilemma), the size of the group and the marginal per capita return are not independent. We define MPCR distance  $d$  as

$$d = \text{MPCR} - 1/N \quad (2)$$

The effect of the MPCR distance on contributions in SPGEs has so far not been discussed in the literature. However, the data obtained in previous experiments already indicate an influence although past papers focus on average behavior and do not refer to first-round contributions specifically. We summarize the evidence below and then go on to formulate the conjectures that we test in our experiment.

For small groups, there is some evidence that the MPCR is very important for the willingness to cooperate. For example, Anna Gunthorsdottir, Daniel Houser, and Kevin McCabe (2007) find that for a fixed group size of four the variation of the MPCR (0.3, 0.5, and 0.75) has a positive but not linear impact on contributions. The MPCR distance for an MPCR of 0.3 here is 0.05. Increasing this distance to 0.25 (MPCR=0.5) has a strong effect, while a further increase to 0.5 (MPCR=0.75) has a rather small effect. These differences between contributions already appear in the first round of the experiment. The authors argue that the increase in contributions

can be explained by the fact that a higher MPCR makes it more effective to invest in the public good (p. 314).<sup>4</sup> But this explanation cannot account for the decreasing strength of the MPCR-effect.

Mark Isaac and James Walker (1988) observe that the impact of an MPCR variation (0.3 vs. 0.75) also depends on the size of the group. While the authors refer to average contributions, the effects are already identifiable in the first rounds. For groups of four the impact of an MPCR variation is much stronger than for groups of ten. This finding can be explained by the fact that for a group of four the MPCR distance is 0.05 for an MPCR of 0.3 and 0.5 for an MPCR of 0.75, while for a group of ten the distance is 0.20 for an MPCR of 0.3 and 0.65 for an MPCR of 0.75. That is, when the MPCR distance is already very high (i.e., 0.2) a further increase of this distance does not much affect behavior, while this is not the case for a low MPCR distance. Another interesting observation made by these authors (p. 191) is that for a high MPCR of 0.75 the group-size effect is rather weak (and not significant) but very strong for a low MPCR of 0.3. Again, this effect might be due to the salience of the dilemma situation. For a high MPCR the MPCR distance is already 0.5 for a group of four (and 0.65 for a group of ten). Apparently, the social dilemma is already fully salient here and, consequently, a further increase has no significant effect on contributions.<sup>5</sup> Mark Isaac, James Walker and Susan Thomas (1984) make a similar observation, which is replicated in Mark Isaac, James Walker and Arlington Williams (1994).<sup>6</sup>

Jeffrey Carpenter (2007) examines MPCRs of 0.375 and 0.75 as well as groups of five and ten subjects. For neither of the two values of the MPCR does he find a significant group-size effect. None of these observations can be explained by either a pure group-size effect or a pure MPCR effect, but are in line with the idea that the salience of the dilemma situation affects first-round behavior. As the lowest MPCR-distance in this experiment is already 0.175, the social dilemma might be close to fully salient so that a further increase does not much affect behavior.

---

<sup>4</sup> In the sense that the total payment to all group members is higher per unit invested in the public good.

<sup>5</sup> That efficiency concerns cannot explain this finding is demonstrated by the following calculation. For an MPCR of 0.3, the group payoff resulting from a one-dollar investment in the public good is \$1.20 in a group of four and \$3.00 in a group of ten. This increase in group size from four to ten significantly increases the average contribution observed in the experiment. For an MPCR of 0.75 the group payoff resulting from a one-dollar investment in the public good is \$3 in a group of four and \$7.5 in a group of ten. In this case, the increase in group size shows no significant impact on the average contribution.

<sup>6</sup> Referring to the data obtained by Mark Isaac, James Walker, and Arlington Williams (1994), Douglas Davis and Charles Holt (1993) mention a different form of interaction between group size and MPCR (compare also Charles Holt and Susan Laury, 2008). They introduce the Minimum Profitable Coalition (MPC) which is the minimum percentage of group members who have to contribute to the public good such that contributing members have a higher payoff as in the case in which no one contributes. They argue that cooperative behavior will be reversely related to the MPC. Our data do not support this idea, however (see our discussion in section 4.2).



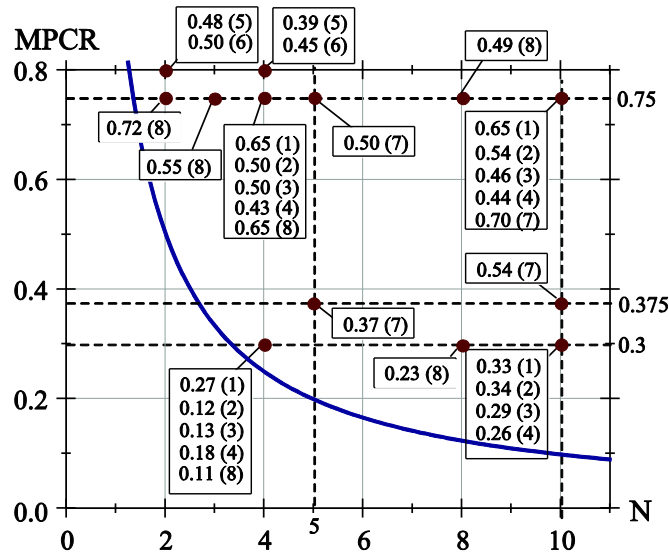


Figure 1: The curve represents, for each group size, the minimum MPCR required to create a public-good situation ( $1/N$ ) and the dots exhibit different combinations of  $N$  and MPCR found in the literature. In the connected boxes we find the respective average contributions. Numbers in parenthesis correspond to the following references: (1) and (2) Isaac et al. (1984), (3) Isaac et al. (1988), (4) Isaac et al. (1994), (5) and (6) Goeree et al. (2002), (7) Carpenter (2007), and (8) Nosenzo et al. (2012).

Figure 1 exhibits the average MPCR and group-size combinations as well as the respective average contribution levels observed in previous SPGEs (for first-round contributions see Figure 2).<sup>7</sup> These observations are plotted against the  $1/N$  curve, which marks the minimal MPCR required to create a public-good situation for the respective group size. In other words, this curve marks for each group size the threshold value of the MPCR such that condition (1\*) is satisfied. Figure 1 suggests the correlation between average contribution levels and MPCR is strong in the neighborhood of the  $1/N$  curve, but becomes weak or non-existent if the *vertical* distance to this curve is larger. For example, for a group of ten and a MPCR of 0.3 the average contribution reported in the literature is about 30 percent. For the same group size and a slightly higher MPCR of 0.375 the reported average contribution increases to 54 percent, while a further doubling of the MPCR has no additional effect in three out of the four reported cases (and in the fourth case the effect is also not significant). To elucidate the relationship between average first-round contributions and MPCR distance, we directly relate them to each other in Figure 2.

<sup>7</sup> See also the overview provided by Daniele Nosenzo, Simone Quercia, and Martin Sefton (2012).

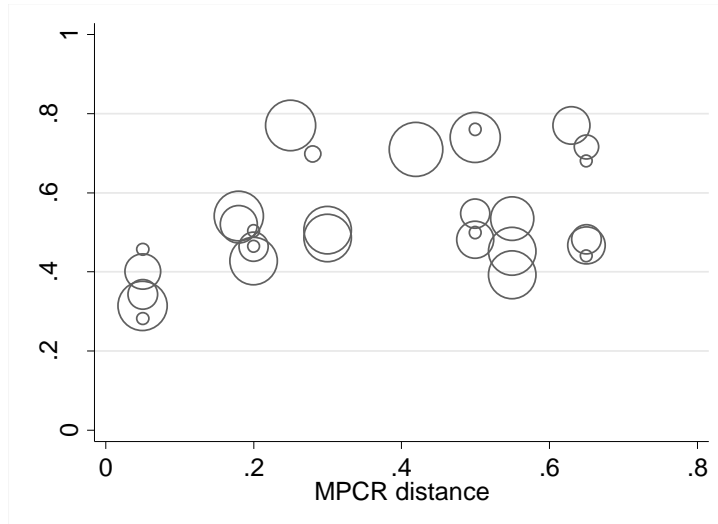


Figure 2: MPCR distance and average first-round contributions (as share of the endowment) reported in the literature as cited in Figure 1<sup>8</sup>

Figure 2 reveals a positive but non-linear relationship and suggests that the MPCR distance  $d$  is likely to be a good indicator for observed first-round contributions (in particular for smaller MPCR distances).

Based on the analysis of previous results summarized above, we argue that  $d$  determines the salience of the social dilemma and, thus, first-round contributions for low levels of  $d$ :

**Conjecture 1:** The first-round contribution  $b_{i1}$  increases with  $d$  for small values of  $d$ .

Fischbacher and Gächter (2010) argue that it is social learning in combination with imperfect conditional cooperation that drives contributions down. Given that imperfect conditional cooperation is based on contribution levels subjects observe after the (first) round, conjecture 1 should also hold for average total contributions and last round contributions.

**Conjecture 2:** The effects of  $d$  on first-round contributions are similar to those of  $d$  on total contributions and on last round contributions.

In case subjects cannot observe (first-)round contributions made by other group members, *social* learning cannot take place. Accordingly, we should observe no or at least a lower decrease of contributions over the course of the experiment.<sup>9</sup>

<sup>8</sup> All data on first-round contributions were kindly provided by the authors or published on personal homepages. The size of the circle represents the number of groups observed under the respective parameter constellation.

<sup>9</sup> There is a large literature on the role of observability in situations in which people have to decide whether to behave selfishly or in some sense other-regarding (Dan Ariely, Anat Bracha, and Stephan Meier, 2009, Roland Benabou and Jean Tirole, 2006.). In particular Mari Rege and Kjetil Telle Kytel (2004) show that implicit social approval is an important driver for contributions in a public good experiment. Contributions increased strongly when they are publicly announced. Therefore the non-observability of individual contributions without feedback may also have a decreasing effect on contributions in general.

**Conjecture 3:** The decrease of contributions is lower without than with feedback.

### 3. Experimental design

In this study we run SPGEs with large groups consisting of 60 and 100 subjects who simultaneously interact under laboratory conditions.<sup>10</sup> For these groups, we consider MPCRs of 0.02, 0.04, and – for groups of 60 – also of 0.06. We use large groups because our conjectures imply that, if groups are large enough, even small changes of very small MPCR values will change the salience of the social dilemma situation. As a side effect, using large groups increases the *external validity* of SPGEs, because real public-good situations (like, e.g., environmental problems) typically affect large groups confronted with very low individual returns from their contributions to the public good.<sup>11</sup>

For control purposes, we additionally run treatments with groups of 30 and 40 with MPCRs of 0.04, 0.06, and 0.12. Given these parameters, the MPCR distances span the range between  $d = 0.007$  and  $d = 0.095$  while the treatments with large groups of 60 and 100 have values of  $d = 0.003$  to  $d = 0.043$ . Moreover, in one of the two treatments with groups of 30 we use a “no-feedback” design, in which subjects decide ten times about their contributions, but never receive information about the contributions made by the other 29 group members in the course of the experiment. Furthermore, we run a control treatment with 20 subjects and an MPCR of 0.06.

Another control involves two treatments with small groups of eight subjects and an MPCR of 0.25. With these treatments we test whether contributions made under a connected-lab design (see below), which we employ to handle large-group SPGEs, differ from those made under a conventional one-lab design.

All treatments were conducted with different subjects to enable between-subject comparisons within our experiment. In all treatments, a standard linear ten-round public-good game was played. The payoff function was identical to equation (1) above, with  $p=1$ . The initial endowment in each of the ten rounds was 120 eurocent. In all treatments (except for control condition C1) we employed a connected-lab design, i.e., groups consisted of subjects who were located

---

<sup>10</sup> As far as we know, 100 is the greatest number of subjects simultaneously participating in a laboratory experiment so far. It is in this sense that our experiment is “non-standard”.

<sup>11</sup> Mark Isaac, James Walker, and Arlington Williams (1994) also investigate large groups, but focus on high MPCRs. Their one observation for a large group with a small MPCR (group size 40 and MPCR=0.03) shows that the average contribution starts at 43 percent, but drops down very quickly to less than 10 percent in round 5 (see Figure 11, p. 28 of their paper). Compared to small groups with a high MPCR this indicates that large groups have more difficulties stabilizing high contribution levels. In their experiment, Mark Isaac, James Walker, and Arlington Williams (1994) neither vary the small MPCR nor the size of large groups faced with a small MPCR. Therefore, the combined effects of both parameters are unclear.

at four different experimental laboratories in Germany (i.e., the laboratories of the Universities of Bonn, Duisburg-Essen, Göttingen, and Magdeburg; see Figure 3) and who simultaneously decided over their individual contribution to the public good.



Figure 3: The location of the four laboratories in Germany

The four laboratories were connected via the Internet. In the treatments with group size 100 (60) (40) (30) (20), on average 25 (15) (10) (7 or 8) (5) subjects participated in each laboratory. All treatments were coordinated by the laboratory in Magdeburg. The communication between laboratories was run via Skype. When entering the respective laboratory, subjects could see a (soundless) video conference of the four laboratories on a computer screen. Thus, each subject had the opportunity to verify that all laboratories were indeed connected and working simultaneously.

At the beginning of each treatment, subjects received written instructions (see Appendix A). Before the start of the first round of the public-good game, they had to answer several questions concerning the payoff rules of the game in order to ensure that they had understood the game correctly. In all treatments (except for the no-feedback treatment) subjects were informed after each round about the amount they had kept, their own contribution, the average contribution to the public good of all group members, their individual payoff from the public good, their individual earnings in the round just completed, and the cumulated earnings over all previous rounds. They knew that the experiment would be finished after ten rounds. Subjects were then paid in cash and left the laboratories. The sessions lasted for about 90 minutes and the average earning was 15.23 euro. The experiment was programmed using zTree (Urs Fischbacher, 2007) and the recruitment of subjects was operated at each of the four locations by ORSEE (Ben Greiner, 2004).

We conducted fifteen treatments and collected data for eight groups (independent observations) per treatment (except for the no-feedback treatment). Table 1 summarizes the treatments. In total 5,220 different subjects participated in the experiment.<sup>12</sup>

Treatment	Group size (average)	MPCR	1/N	MPCR distance	Sessions / indep. obs.	Lab
T1 (8-0.25)	8 (8)	0.25	0.125	0.125	8	connected
T2 (30-0.04)	30 (30)	0.04	0.033	0.007	8	connected
T3 (30-0.06)	30 (30)	0.06	0.033	0.027	8	connected
T4 (30-0.12)	30 (30)	0.12	0.033	0.087	8	connected
T5 (40-0.04)	40 (40)	0.04	0.025	0.015	8	connected
T6 (40-0.06)	40 (40)	0.06	0.025	0.035	8	connected
T7 (40-0.12)	40 (40)	0.12	0.025	0.095	8	connected
T8 (60-0.02)	60 (60)	0.02	0.017	0.003	8	connected
T9 (60-0.04)	60 (60)	0.04	0.017	0.023	8	connected
T10 (60-0.06)	60 (60)	0.06	0.017	0.043	8	connected
T11 (100-0.02)	100 (99.6)	0.02	0.010	0.010	8	connected
T12 (100-0.04)	100 (95.4)	0.04	0.010	0.030	8	connected
C1 (8-0.25L)	8 (8)	0.25	0.125	0.125	8/32	local
C3 (30-0.06nf)	30 (30)	0.06	0.033	0.027	1/30	connected
C11 (20-0.06)	20 (20)	0.06	0.050	0.010	8	connected

Table 1: Treatments

Across twelve of the treatments we systematically vary group size and MPCR (T1 to T12) while three additional treatments serve as controls (C1, C3, and C11). The first treatment T1 with a small group size of eight and an MPCR of 0.25 serves as the baseline with the highest MPCR distance. In T2 (30-0.04), T3 (30-0.06), and T4 (30-0.12) groups of 30 subjects played the public-good game facing MPCRs of 0.04, 0.06, and 0.12, respectively. Treatments T5 (40-0.04), T6 (40-0.06), and T7 (40-0.12) were conducted with a group size of 40 and the same MPCRs. For the larger groups we also considered a small MPCR of 0.02. In Treatments T8 (60-0.02), T9 (60-0.04), and T10 (60-0.06) groups of 60 subjects played the game with MPCRs of 0.02,

<sup>12</sup> As far as we know this means that our project is also one of the biggest experimental investigations conducted in economics so far.

0.04, and 0.06. In Treatments T11 (100-0.02) and T12 (100-0.04) groups of 100 faced MPCRs of 0.02 and 0.04. Because of no-shows, the number of 100 subjects per group was not always reached. In T11 and T12 the average numbers of subjects were 99.6 and 95.4.<sup>13</sup>

In all treatments except for C1, subjects were distributed over the four laboratories. To test whether the connected-lab design had any influence on subjects' decisions we also ran the control treatment C1. In C1 (8-0.25L), we used an MPCR of 0.25 and local groups of eight subjects; that is, each laboratory locally collected data of eight independent groups each consisting of eight subjects. This treatment can directly be compared to T1 (8-0.25) that employed the same parameter constellation, but was conducted with subjects in each of the four laboratories that formed a group of eight in the connected lab.

Across treatments subjects were informed about the outcome of the game after each round. In C3 (30-0.06nf), however, we did not provide feedback. Other than this, the treatment was identical to T3 (30-0.06). Note that in C3 individual contributions are independent of each other. Therefore one group consisting of 30 subjects yields 30 independent observations.

In the third control treatment C11 (20-0.06) we introduce a new parameter constellation. In this treatment groups of 20 subjects played the public-good game facing an MPCR of 0.06. This treatment can be compared to T11 (100-0.02). The parameters of both treatments yield the same MPCR distance of 0.01. If conjectures 1 and 2 are correct, we should observe very similar contributions in the two treatments although they differ a lot with respect to group size and MPCR.

The combinations of group-sizes and MPCRs employed in T2 to T12 and C11 are displayed in Figure 4. This figure spans the  $N$ -MPCR space for  $N$  between zero and 100. In addition to the eleven points representing our treatments, it shows the  $1/N$  curve, which for each group size provides the minimal MPCR for the public-good condition to be satisfied. We see, for example, that an increase in the group size from 60 to 100 results in a smaller change of the MPCR distance than the increase of the MPCR from 0.02 to 0.04. It follows from our conjectures that, for these treatments, the MPCR effect should be stronger than the group size effect.

---

<sup>13</sup> Since each individual decision in the first round is an independent observation, we can check if groups of less than (but close to) 100 behaved differently from those with exactly 100 subjects. We found no significant difference ( $p > 0.390$  two-sided Mann-Whitney- $U$  tests).

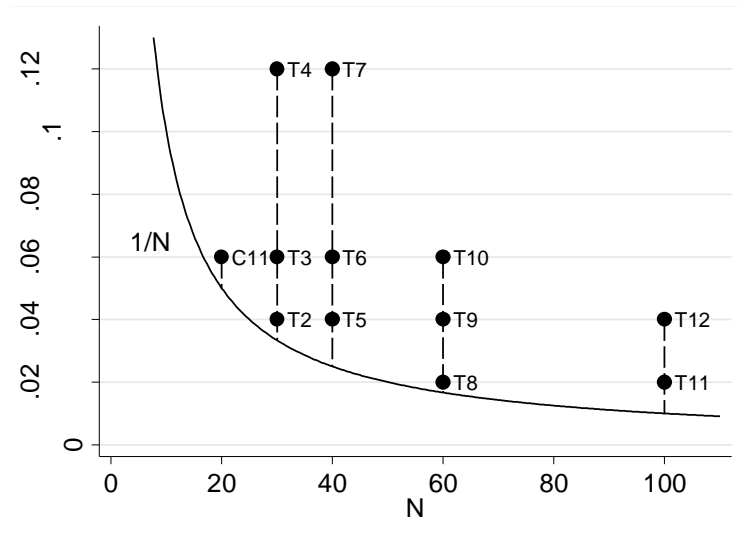


Figure 4: MPCR, group size, and MPCR distance (T2 to T12 and C11)

## 4. Results

### 4.1 Overview

Before analyzing behavior in more detail, we have to check the critical assumption that the connected-lab design that connects group members at different laboratories in Germany has no influence on subjects' behavior. Comparing T1 (8-0.25) and C1 (8-0.25L), we cannot reject the hypothesis that average contributions are independent of whether we conduct the experiment in one laboratory or in a connected laboratory.<sup>14</sup> There is neither a significant difference between local groups and groups in the connected lab regarding average contributions made over all ten rounds nor regarding average contributions made in each of the rounds. From a methodological point of view this finding is good news, because it appears that the capacity of laboratories can be multiplied by virtually connecting them without inducing significant behavioral effects. Figure 5 illustrates the average contributions over the ten rounds in T1 and C1.

<sup>14</sup>  $p > 0.264$  for comparing contributions in each of the 10 rounds and  $p = 0.685$  for comparing overall contributions (two-sided Mann-Whitney  $U$  tests; unless reported otherwise, in the following all  $p$ -value refer to two-sided Mann-Whitney  $U$  tests). All data are available at [http://www.ovgu.de/vwl3/experiments/pg\\_large\\_groups/](http://www.ovgu.de/vwl3/experiments/pg_large_groups/). A summary of descriptive statistics is included in Appendix B.

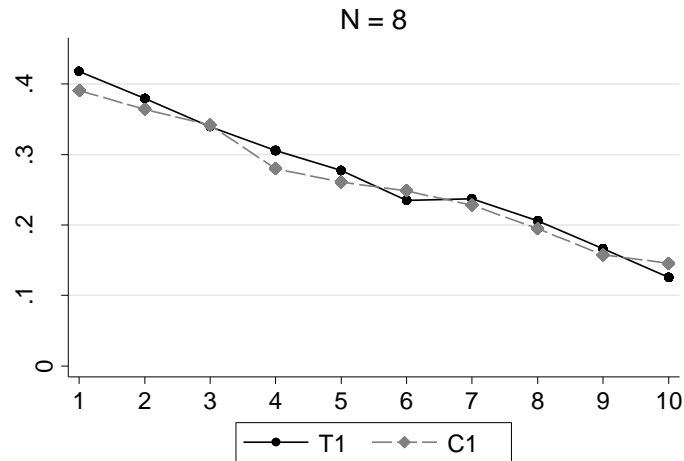


Figure 5: Mean contributions in treatments T1 and C1

Figure 6 shows the average share of contributions across the eleven remaining treatments over the 10 rounds sorted by group size. Figure 7 presents the contributions in the two other control treatments. In all of these treatments we observe the typical decay of contributions over time. Table 2 summarizes the descriptive statistics of all treatments including participants' age and gender.

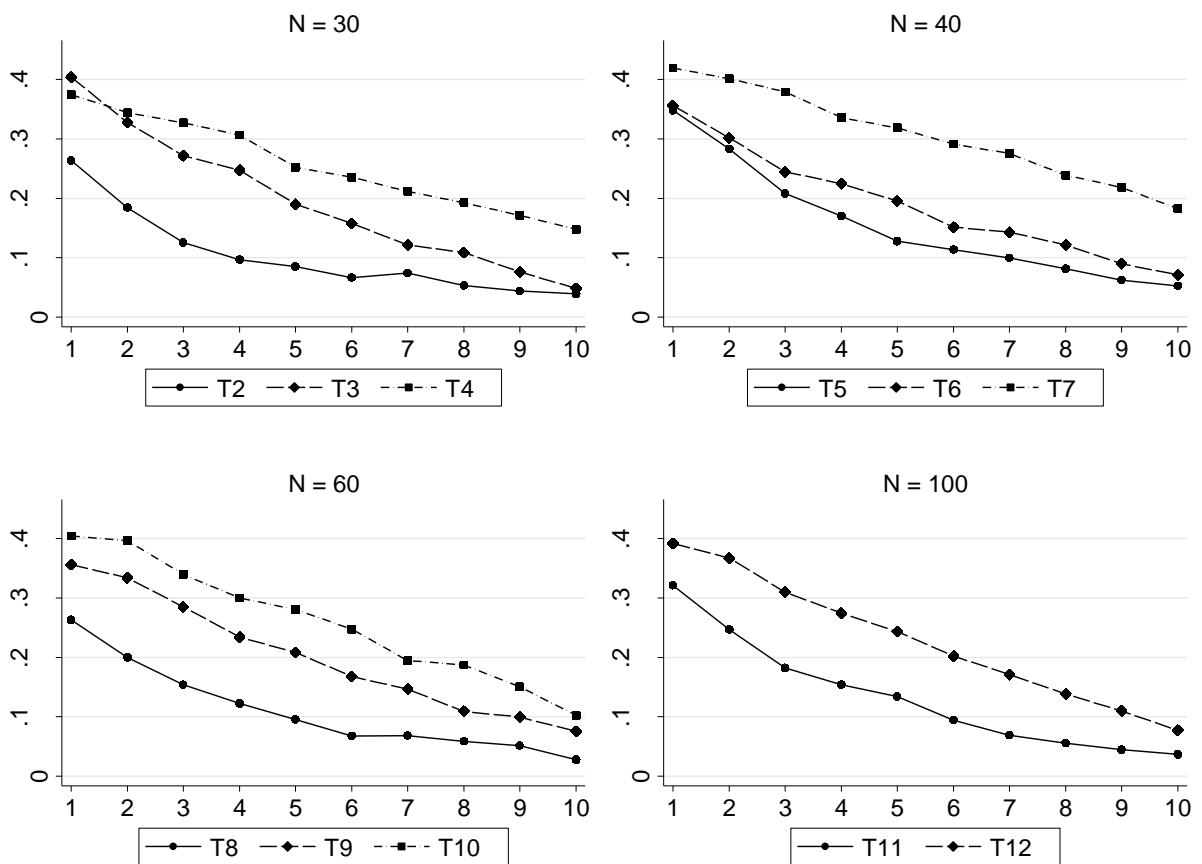


Figure 6: Mean contributions across treatments T2 to T12 over group size



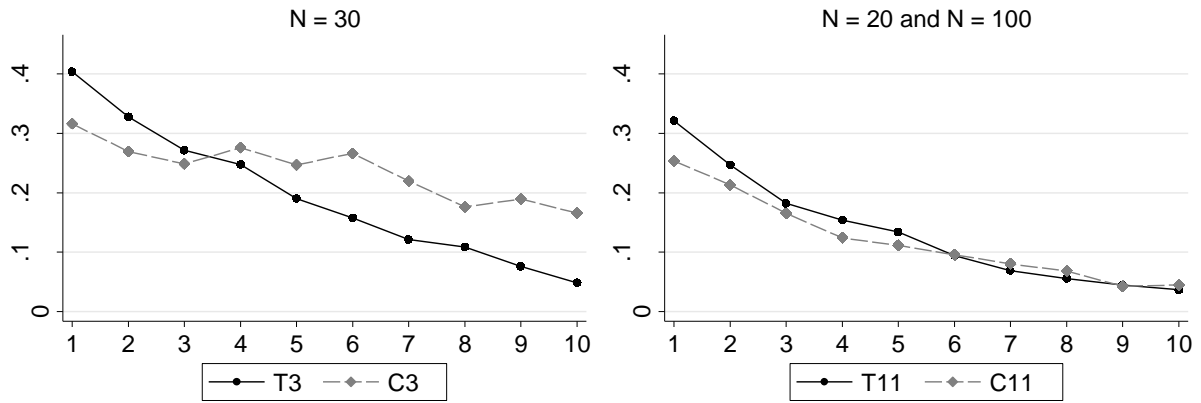


Figure 7: Mean contributions in the control treatments C3 and C11

Treatment	Contributions			Age in years	Female dummy
	Round 1	Round 10	All rounds		
T1 (8-0.25)	0.418 (0.127)	0.126 (0.101)	0.269 (0.110)	23.297 (3.077)	0.547 (0.221)
T2 (30-0.04)	0.264 (0.017)	0.040 (0.032)	0.103 (0.026)	23.287 (0.842)	0.471 (0.072)
T3 (30-0.06)	0.404 (0.060)	0.049 (0.020)	0.195 (0.029)	22.729 (0.421)	0.525 (0.131)
T4 (30-0.12)	0.374 (0.065)	0.148 (0.028)	0.256 (0.039)	23.283 (0.721)	0.525 (0.087)
T5 (40-0.04)	0.348 (0.074)	0.052 (0.034)	0.155 (0.038)	22.734 (0.660)	0.531 (0.084)
T6 (40-0.06)	0.356 (0.036)	0.071 (0.029)	0.190 (0.028)	23.022 (0.627)	0.478 (0.093)
T7 (40-0.12)	0.419 (0.055)	0.183 (0.06)	0.306 (0.059)	22.784 (0.508)	0.569 (0.079)
T8 (60-0.02)	0.263 (0.047)	0.028 (0.013)	0.111 (0.021)	22.979 (0.462)	0.519 (0.049)
T9 (60-0.04)	0.356 (0.028)	0.075 (0.030)	0.202 (0.019)	23.125 (0.502)	0.517 (0.076)
T10 (60-0.06)	0.404 (0.054)	0.103 (0.031)	0.260 (0.046)	22.723 (0.545)	0.494 (0.076)
T11 (100-0.02)	0.321 (0.047)	0.037 (0.009)	0.134 (0.021)	23.709 (0.599)	0.439 (0.034)
T12 (100-0.04)	0.391 (0.053)	0.077 (0.008)	0.228 (0.031)	23.169 (0.539)	0.487 (0.083)
C1 (8-0.25L)	0.391 (0.13)	0.146 (0.127)	0.261 (0.146)	22.773 (1.322)	0.590 (0.188)
C3 (30-0.06nf)	0.316 (0.32)	0.166 (0.285)	0.238 (0.265)	24.067 (3.403)	0.400 (0.498)
C11 (20-0.06)	0.253 (0.079)	0.045 (0.035)	0.120 (0.053)	22.581 (0.463)	0.500 (0.080)

Table 2: Summary of means and standard deviations (in brackets)

Given the descriptive statistics of our results some remarks are in order. First, notice that the MPCR effect for a given group size displayed in Figure 6 is rather strong. On the other hand, the differences between groups of different sizes but the same MPCR (for example T2 versus T5, T3 versus T6) seem to be much smaller. The comparison of T11 and C11 shows that indeed the contributions in the two treatments are very similar over the whole course of the experiment, although group size (20 versus 100) and MPCR (0.06 versus 0.02) differ a lot. Remember that the MPCR distance is the same in both treatments. These observations might be regarded as first evidence that our conjectures are confirmed by our experimental results. In the next subsection we present the results of regression analyses that control for age, gender, and location of participants.

#### 4.2 MPCR distance and contributions

Conjecture 1 states that first-round contributions are an increasing function of the MPCR distance. Furthermore, it states that the effect is most pronounced for small values of  $d$ . This can be captured by assuming that the marginal effect of the MPCR distance  $d$  is decreasing. More formally, for the following analysis we hypothesize that  $b_{i1}$  is a function that is increasing but concave in  $d$ :

$$\frac{\partial b_{i1}(d, \cdot)}{\partial d} > 0, \frac{\partial^2 b_{i1}(d, \cdot)}{\partial^2 d} < 0 \quad (3)$$

In our econometric estimation we model an individual's contribution  $b_{i1}$  as a linear function of individual specific control variables  $x_i$  consisting of age, gender and location dummies.<sup>15</sup> The effect of MPCR distance is captured by a linear term  $d_i$  as well as quadratic term  $d_i^2$ .  $\varepsilon_i$  denotes the error-term.

$$b_{i1} = \beta_0 + \beta_1 d_i + \beta_2 d_i^2 + \beta_3' x_i + \varepsilon_i \quad (4)$$

Tables 3a to 3c present several regression results. Table 3a includes the above specification in column 3 as well as specifications that either ignore the linear term (columns 2) or the quadratic term (columns 1). Treatments T1 (8-0.25), T4 (30-0.12) and T7 (40-0.12) are characterized by rather high MPCR distances of 0.125, 0.087 and 0.095. Therefore, we conduct the regressions excluding treatments T1, T4 and T7 with  $d \leq 0.043$  (on the left side of the table) and including treatments T1, T4, and T7 (on the right side). We also vary the dependent variable. The above specification focuses on Conjecture 1. To check whether the influence of the MPCR distance

---

<sup>15</sup> We do not consider the effect of these controls in detail in this study. Roughly: Female (older) subjects contribute more than male (younger) subjects, and subjects in Essen and Göttingen contribute more than subjects in Bonn and Magdeburg.

persists throughout the ten rounds of the game as stated in Conjecture 2, we also run the regression on average contributions  $b_i$  and last-round contributions  $b_{i10}$  (Tables 3b and 3c).

	$d \leq 0.043$			Full Sample		
	(1)	(2)	(3)	(1)	(2)	(3)
MPCR Distance	3.387*** (0.436)		7.493*** (1.654)	1.247*** (0.203)		4.555*** (0.649)
(MPCR Distance) <sup>2</sup>		65.302*** (10.193)	-91.193** (36.160)		8.249** (1.788)	-31.626*** (5.992)
<i>N</i>	4,120	4,120	4,120	4,744	4,744	4,744
<i>N</i> left-censored	763	763	763	839	839	839
<i>N</i> right-censored	367	367	367	443	443	443
AIC	5536.468	5553.130	5530.629	6421.643	6450.398	6384.310

Table 3a – Tobit regressions on contributions in round 1

	$d \leq 0.043$			Full Sample		
	(1)	(2)	(3)	(1)	(2)	(3)
MPCR Distance	3.448*** (0.283)		4.892*** (1.033)	1.632*** (0.157)		4.547*** (0.418)
(MPCR Distance) <sup>2</sup>		70.081*** (7.172)	-32.054 (24.595)		11.927*** (1.558)	-27.838*** (4.055)
<i>N</i>	2760	2760	4,120	4,744	4,744	4,744
<i>N</i> left-censored	355	355	501	540	540	540
<i>N</i> right-censored	5	5	11	15	15	15
AIC	-1853.508	-1806.391	-1856.351	-1617.257	-1455.968	-1751.084

Table 3b – Tobit regressions on average contributions over all rounds

	$d \leq 0.043$			Full Sample		
	(1)	(2)	(3)	(1)	(2)	(3)
MPCR Distance	1.353*** (0.172)		1.344** (0.662)	1.005*** (0.097)		2.027*** (0.277)
(MPCR Distance) <sup>2</sup>		27.903*** (3.890)	0.194 (15.467)		7.973*** (1.029)	-9.578*** (2.611)
<i>N</i>	4,120	4,120	4,120	4,744	4,744	4,744
<i>N</i> left-censored	2,431	2,431	2,431	2,689	2,689	2,689
<i>N</i> right-censored	55	55	55	80	80	80
AIC	3387.769	3392.54	3389.769	4227.447	4274.144	4211.691

This table reports average marginal effects on the observed contributions. Regressions include controls for age, gender and location. Standard errors clustered on the session level are given in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-sided).

Table 3c – Tobit regressions on contributions in round 10

Let us first consider first-round contributions. As the regression results in Table 3a show, the full model yields the best fit in both, the restricted and the full sample as indicated by Akaike's information criterion (AIC). The results are illustrated by Figure 8a, which includes average first-round contributions as well as the predictions of our three model specifications.

What conclusions can we draw from these results? On the small dataset with  $d \leq 0.043$  the increase in  $d$  from 0.01 to 0.02 yields a predicted increase in first-round contributions by 4.8 percent while the increase from 0.03 to 0.04 yields a predicted increase of 1.1 percent. On the full data set the increase in  $d$  from 0.01 to 0.02 yields a predicted increase of 3.6 percent while the increase from 0.03 to 0.04 yields a predicted increase of 2.3 percent. In line with Conjecture 1,  $d$  exhibits a positive, but diminishing influence on first-round contributions for small values of  $d$ . In fact, the predicted influence of  $d$  becomes negative for large values of  $d$ . This is, of course, an artifact of our model specification. As the results of previous studies summarized in section 2 and displayed in Figure 2 suggest, the level of first-round contributions in studies with higher MPCR distances ranging up to 0.65 never falls considerably below the level of contributions we observe for  $d=0.125$  in T2.

Do the differences we observe in the first-round carry over to the remaining rounds? In Conjecture 2 we propose that the influence of the MPCR distance persists when considering average contributions over all rounds or last-round contributions. Figure 8b displays average contributions made over all rounds over the MPCR distance  $d$ . Again, we observe a positive, but diminishing increase of contributions in  $d$ . This is confirmed by the regressions reported in Table 3b. The effect of  $d$  is significant and positive. And, as for first round-contributions, the full model fits best for both data sets based on the AIC because it captures the diminishing effect of  $d$  on contributions. In the last round, contributions have considerably decayed and are low across treatments. But in this round the positive relationship between  $d$  and contributions still persists. In this case, for the low  $d$  treatments the linear model captures the relationship best. Once we include T1, T4, and T7, it is again the full model that yields the best fit because the influence of  $d$  diminishes.

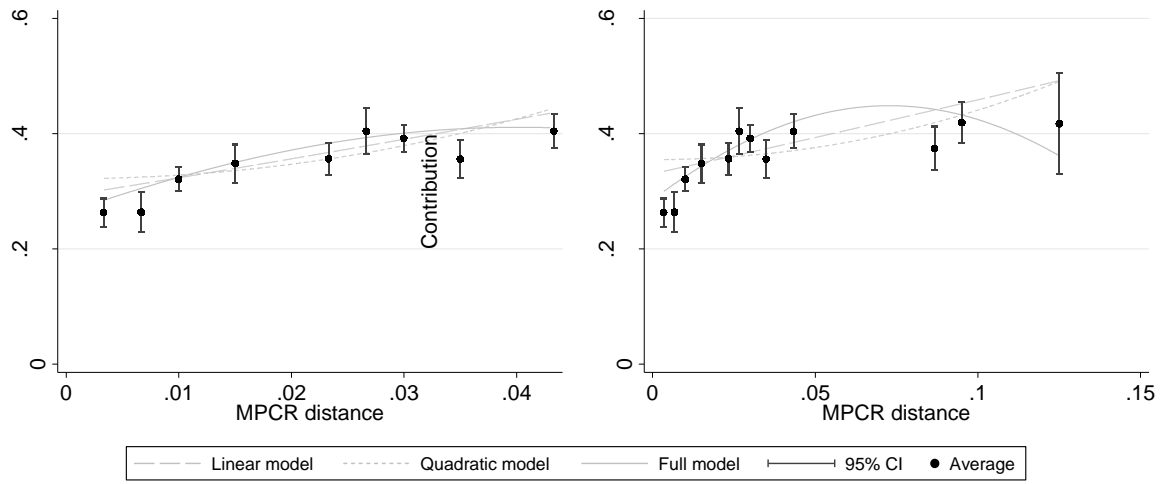


Figure 8a: Average contributions and MPCR distance in round 1.

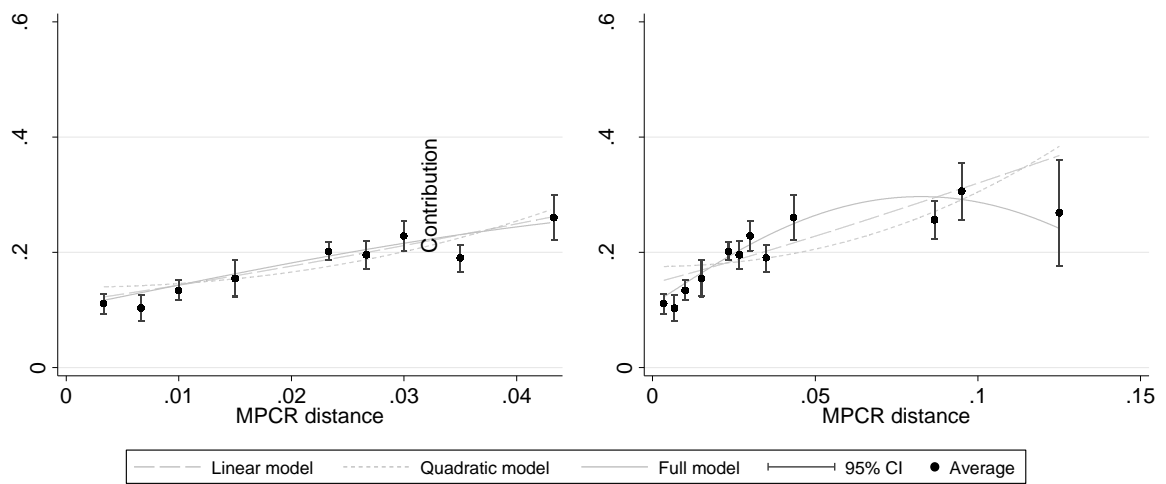


Figure 8b: Average contributions and MPCR distance over all rounds

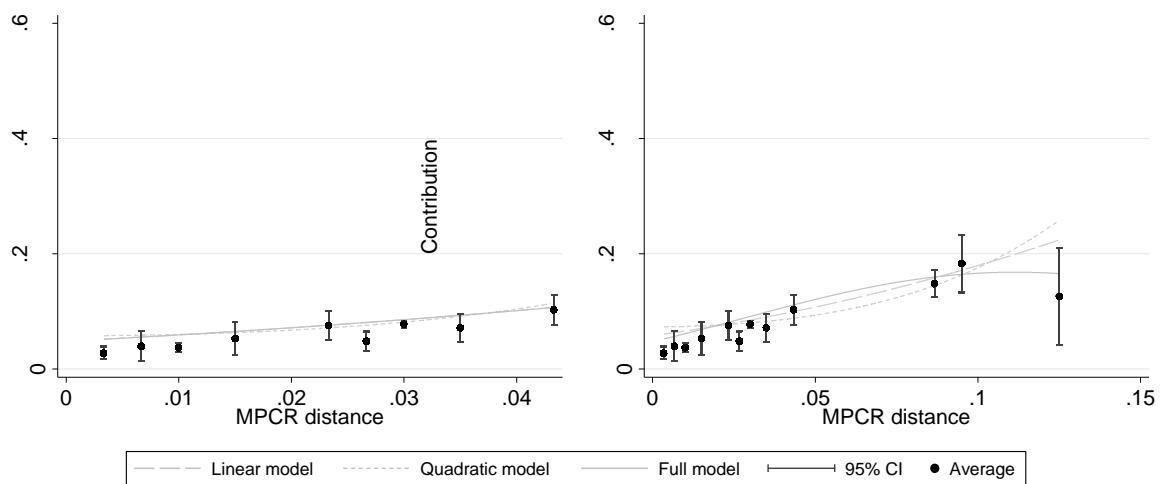


Figure 8c: Average contributions and MPCR distance in round 10

Douglas Davis and Charles Holt (1993) mention an alternative form of interaction between group size and the MPCR. They argue that cooperative behavior is inversely related to the size of the Minimum Profitable Coalition (MPC). The MPC is the minimum percentage of group members who have to fully contribute to the public good such that their resulting individual payoff is higher than the individual payoff, which results in case that nobody contributes. To compare the MPC and the MPCR distance we run regressions with both variables and compare the resulting AIC. We also run regressions with both variables in one model and compare the significance. We find that the MPCR distance outperforms the MPC in all regressions suggesting that the former has more predictive power than the latter (all regressions are included in Appendix B).

#### 4.3 Dynamics of contributions

As Figure 6 has shown, average behavioral patterns observed for groups consisting of 8, 30, 40, 60 or 100 subjects are very similar to each other. Not only are the average first-round contributions positive across treatments, but there is also a decay of average contributions over the ten rounds. In particular, comparing average first-round contributions with those made in the last round, we find significant differences for all treatments ( $p \leq 0.012$ , Wilcoxon tests, two-sided). Moreover, in all treatments average last-round contributions are significantly lower than total contributions. However, when comparing average first-round contributions with total contributions, we find significant differences for all treatments ( $p \leq 0.012$ ) except for C3. Only in the treatment without feedback average total contributions do not differ significantly from first-round contributions ( $p = 0.452$ ).

This can be regarded as evidence that imperfect conditional cooperation and social learning takes place in large groups as well. C3 only differs from T3 with respect to the information subjects receive during the course of the game. In both treatments we observe a decay of contributions. To check whether the strength of the decay depends on feedback we run the following panel regression. Let  $b_{it}$  denote the contribution of player  $i$  in round  $t$  of the game. We model the contribution as a linear function of round  $t$  and the treatment characteristics as indicated by a vector of treatment dummies  $T_i$  as follows

$$b_{it} = \beta_0 + \beta_1' T_i + \beta_2 t_{it} + \beta_3' T_{itit} + \beta_4' x_i + \eta_i + \varepsilon_{it}.$$

We assume independent and normally distributed error terms denoted by  $\eta_i$  and  $\varepsilon_{ij}$  that capture differences between individuals and idiosyncratic effects within rounds. The vector of dummy variables  $T_i$  indicates individual's  $i$  respective treatment (T1 to T12, C1, C11) while C3 serves

as the baseline. This regression reveals a significant and negative effect of the round on contributions in C3 ( $\beta_2=-0.025$ ,  $p=0.000$ ). However, the interactions of the treatment dummies and the round reveal that the decay is significantly stronger in Treatments T1 to T12 than in C3 ( $\beta_{3i}\leq-0.010$ ,  $p\leq 0.024$ ). In treatment T7 this effect is only weakly significant ( $p=0.067$ ). It is not significant in treatment T4 ( $\beta_{34}=-0.009$ ,  $p=0.115$ ). See Appendix C for the full regression results.

The significant difference in decay between C3 and T3 supports Conjecture 3 that stated less decay in contributions without feedback than with feedback. It appears that social learning in the feedback treatment drives contributions down over all 10 rounds, while the absence of social learning possibilities stabilizes contributions in the no-feedback treatment over several rounds. The restriction that subjects cannot observe each other's contributions in the no-feedback treatment might have also reduced their initial willingness to cooperate. Contributions are significantly lower in the first round of C3 than in that of T3 ( $p=0.001$ , Mann-Whitney- $U$  test, two-sided). This observation might be interpreted as support for previous results of experiments showing an impact of implicit social approval (e.g., Dan Ariely, Anat Bracha, and Stephan Meier, 2009, see footnote 11) in the context of large groups.

## **5. Discussion**

It is known from the literature that both group size and marginal per capita return (MPCR) of an investment in the public good influence contributions in standard public-good experiments. But both effects interact with each other. The main contribution of the present paper is that it presents and tests an explanation for this interaction. Our explanation is derived from the idea that the salience of a social dilemma determines first-round contributions. The higher the salience of the dilemma, the more subjects are aware of this situation and the more subjects can be sure that everybody else knows that contributing to the public good will make all group members better off. If the MPCR is larger than  $1/N$  (given that the payoff from the private investment is normalized to 1), it is efficient to invest in the public good. We argue that the dilemma situation is salient to subjects only when the distance between the MPCR and  $1/N$  is high enough. The interaction of group size and MPCR is expressed in this distance that we label "MPCR distance". According to our explanation, the MPCR distance determines first-round contributions in a particular way: As long as it is small (i.e., as long as the dilemma situation is not fully salient to group members), an increase will have a strong and positive effect on first-round contributions, because it makes the social dilemma more salient.

To test our hypothesis, we designed the experiment in such a way that small changes in small MPCR distances can be investigated. We did this by conducting experiments with large groups and very small MPCRs. Both are common features of real-world public good problems. In this environment even very small changes of the MPCR should strongly affect contributions. This is exactly the effect we observe. Furthermore, we find that the MPCR distance also explains considerable variation of contributions between treatments in subsequent rounds.<sup>16</sup>

Our MPCR distance explanation complements the theory by Fischbacher and Gächter (2010) by explaining variations in first-round contributions. Fischbacher and Gächter (2010) focus on the dynamics of cooperation and identify social learning as the driving force for the decay of contributions in repeated public-good games. Our study provides further support for this explanation by demonstrating that social learning drives the decay of contributions also in large groups.

We make two additional contributions. First, we introduce a connected-laboratory design. By connecting four laboratories we are able to conduct experiments with large groups while maintaining the control of a laboratory environment. Second, using large groups and small MPCRs we increase the external validity of standard public-good experiments. In the real world, public good problems usually involve large groups and these large groups are very often confronted with small MPCRs (e.g., in environmental settings or voting).

Our results have an important implication for the experimental investigation of public good situations because they demonstrate that behavioral dynamics are the same in small groups with a high MPCR and in large groups with a small MPCR. Thus small groups seem to be well suited to cover essential characteristics of public goods in a laboratory situation. At the same time, our explanation of the behavior in those experiments implies that for the investigation of real public good problems the salience of the social dilemma seems to be of great importance. Therefore, our results have several practical implications. They demonstrate that in order to motivate people to contribute to a public good, the salience of the dilemma is crucial. For example, in the case of environmental problems it is important that, first, people know that their own cooperative contribution is efficiency enhancing, and that, second, they are convinced that the social dilemma situation is common knowledge for all people in the group. Probably, the salience of

---

<sup>16</sup> This effect is also in line with an explanation involving reference points that are initially set by the MPCR distance. A reference point may be an anchor for the decision of how much to contribute or it might be an element of the “social situation”, which prescribes the behavior necessary to maintain the social identity of a person. See, e.g., the model by George Akerlof and Rachel Kranton (2000), which is based on the assumption that social identity is an important argument in utility functions. See also the psychological and sociological literature on the importance of social identity cited therein.



real world public good problems will not only depend on the MPCR distance. Identifying additional influences is an important task for future research.

## References

- Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review*, 99(1): 544–55.
- Akerlof, George A. and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics*, 115(3): 715–53.
- Andreoni, James. 1988. "Why free ride? Strategies and Learning in Public Goods Experiments." *Journal of Public Economics*, 37: 291–304.
- Andreoni, James. 1995. "Cooperation in Public-Goods Experiments. Kindness or Confusion?" *American Economic Review*, 85(4): 891–904.
- Beckenkamp Martin, Christoph Engel, Andreas Glöckner, Bernd Irlenbusch, Heike Hennig-Schmidt, Sebastian Kube, Michael Kurschilgen, Alexander Morell, Andreas Nicklisch, and Hans-Theo Normann. 2014. "First Impressions are More Important than Early Intervention: Qualifying Broken Windows Theory in the Lab." *International Review of Law and Economics*, 37: 126–136.
- Benabou, Roland and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review*, 96(5):1652–1678.
- Burlando, Roberto and Francesco Guala. 2005. "Heterogeneous Agents in Public Good Experiments." *Experimental Economics*, 8: 35–54.
- Burlando, Roberto and John D. Hey. 1997. "Do Anglo-Saxons Free-ride More?" *Journal of Public Economics*, 64(1): 41–60.
- Carpenter, Jeffrey P. 2007. "Punishing Free-riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods." *Games and Economic Behavior*, 60(1): 31–51.
- Chaudhuri, Annanish. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature." *Experimental Economics*, 14(1): 47–83.
- Croson, Rachel. 2007. "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Good Games." *Economic Inquiry*, 45(2): 199–216.
- Davis, Douglas D. and Charles A. Holt. 1993. Public Goods, Externalities, and Voting. In Douglas D. Davis and Charles A. Holt (Eds.): *Experimental Economics*, 317–379. Princeton, N J: Princeton University Press.

- Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10(2): 171–178.
- Fischbacher, Urs and Simon Gächter. 2010. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments." *American Economic Review*, 100(1): 541–56.
- Fischbacher, Urs, Simon Gächter, and Ernst Fehr. 2001. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters*, 71(3): 397–404.
- Gächter, Simon and Ernst Fehr. 1999. "Collective Action as a Social Exchange." *Journal of Economic Behavior & Organization*, 39: 341–69.
- Gächter, Simon and Christian Thöni. 2005. "Social Learning and Voluntary Cooperation among Like-Minded People." *Journal of the European Economic Association*, 3(2-3): 303–14.
- Goeree, Jacob K., Charles A. Holt, and Susan K. Laury. 2002. "Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior." *Journal of Public Economics*, 83(2): 255–276.
- Greiner, Ben. 2004. "The Online Recruitment System ORSEE 2.0-A Guide for the Organization of Experiments in Economics." In: Kurt Kremer, Volker Macho (Eds.): *Forschung und wissenschaftliches Rechnen. GWDG Bericht 63. Ges. für Wiss. Datenverarbeitung*, 79–93, Göttingen.
- Gunnthorsdottir, Anna, Daniel Houser, and Kevin McCabe. 2007. "Disposition, History and Contributions in Public Goods Experiments." *Journal of Economic Behavior & Organization*, 62(2): 304–15.
- Holt, Charles A. and Susan K. Laury. 2008. "Theoretical Explanations of Treatment Effects in Voluntary Contributions Experiments." In Charles R. Plott and Vernon L. Smith (Eds.): *Handbook of Experimental Economics Results, Volume 1*, 846–855. New York: Elsevier Press.
- Isaac, R. Mark, James M. Walker, and Susan H. Thomas. 1984. "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." *Public Choice*, 43(2): 113–49.
- Isaac, R. Mark and James M. Walker. 1988. "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism." *Quarterly Journal of Economics*, 103(1): 179–99.

- Isaac, R. Mark, James M. Walker, and Arlington W. Williams. 1994. "Group Size and the Voluntary Provision of Public Goods. Experimental Evidence Utilizing Large Groups." *Journal of Public Economics*, 54(1): 1–36.
- Keser, Claudia and Frans van Winden. 2000. "Conditional Cooperation and Voluntary Contributions to Public Goods." *Scandinavian Journal of Economics*, 102(1): 23–39.
- Ledyard, John O. 1995. "Public Goods: A Survey of Experimental Research." In John H. Kagel, A. E. Roth (Eds.): *The Handbook of Experimental Economics*, 111–94. Princeton, N.J: Princeton University Press.
- Nosenzo, Daniele, Simone Quercia and Martin Sefton. 2012. "Cooperation in Small Groups: The Effect of Group Size." Forthcoming: *Experimental Economics* (DOI 10.1007/s10683-013-9382-8).
- Ockenfels, Axel and Joachim Weimann. 1999. "Types and Patterns: An Experimental East-West-German Comparison of Cooperation and Solidarity." *Journal of Public Economics*, 71: 275–87.
- Palfrey, Thomas R. and Jeffrey E. Prisbrey. 1997. "Anomalous Behavior in Public Goods Experiments. How Much and Why?" *American Economic Review*, 87: 829–46.
- Rege, Mari and Kjetil Telle, 2004. "The Impact of Social Approval and Framing on Cooperation in Public Good Situations." *Journal of Public Economics*, 88: 1625–1644.
- Weimann, Joachim. 1994. "Individual Behavior in a Free Riding Experiment." *Journal of Public Economics*, 54: 185–200.

## Appendix A: Instructions Treatment 5 (100-0.02)

### Instructions 100

Preliminary: You are participating in an economic experiment focusing on decision making. If there are any questions left after having read these instructions or during the experiment, please raise your hand. We will then come to your cubicle.

While participating in the experiment, you have to take a sequence of decisions. You will earn money. But, how much money you will earn will depend both on your decision and the decisions of the other participants. Your total earnings will be paid in cash at the end of the experiment. Both your decisions and your payoff are confidential, i. e. no other participant will receive this information..

You are part of a group of 100 participants. These 100 people are located in four experimental laboratories across Germany, connected via internet. All group members received the same instructions. Moreover, the laboratories are connected via a video connection. If you have any doubts about this procedure, please take a look at our video conference!

You and the other 99 group members are facing the following identical decision situation during 10 consecutive rounds: In each round, you receive an endowment of 120 Euro Cent. You decide how much of this endowment you want to “keep”, and how much you want to “contribute”. Each contribution  $x$  is creating an amount  $0.02 \cdot x$  for *each* group member (including the contributor). That means that for every Euro Cent you contribute, the whole group will be paid 2 Euro Cent ( $0.02 \cdot x \cdot 100$ ). For each Euro Cent you do contribute, you will be paid 0.02 Euro Cent like all other group members. That part of your endowment that you do not contribute (i. e. that you “keep”), you keep for yourself.

Summing up in one formula, your earnings in Euro Cent per round are as follows:

$$120 - \text{Your Contribution} + 0.02 \cdot (\text{Sum of all group members' contributions})$$

Please note that your contribution per round can be any amount between 0 and 120 Euro Cent and that all group members are facing an identical decision situation. After each round you will be informed on the amount you kept, your contribution, the average contribution of *all* 100 group members, your payoff based on the contributions of all group members, your payoff in

the respective round and your payoff cumulated over all periods. Moreover, you will see a table listing the same information for all previous rounds.

*Practice rounds:* Before starting with the experiment, you have the opportunity to decide in three practice rounds. In these practice periods, the average contribution of all other group members will be *given* as it is randomly generated. Furthermore, your own contribution will be preset, too. Your task is to calculate the earnings in the respective round yourself. To that end, we provide you with a calculator, paper, and pencil. After having entered your solution into the respective box, please click on the “Solution” button. You then will be informed on whether your answer is right or wrong. Also the calculation method will be shown. If you have any questions during the practice rounds, please raise your hand. Right after the practice periods are over, the experiment will start automatically.

*Payoff:* Please stay in your cubicle after all 10 rounds have ended. You will be called individually to receive your payoff. Please hand in your participation number (which you have drawn at the beginning of the experiment) and enter your name and signature in the payment list. Please leave the laboratory after having received your money.

Finally we would like to ask you to not talk to anybody about the content of this experiment to avoid influencing future participants. Thank you for your cooperation!

## Appendix B: Comparison of MPC and MPCR distance

The MPC explanation (Davis and Holt, 1993) provides another explanation that builds on a simple relationship of MPCR and group size  $N$ . Our treatments also allow testing it.<sup>17</sup> Figure B1 depicts our treatments together with iso-MPC curves for MPC values of 27.8 (T4, T10), 41.7 (T6, T9), 50.0 (T1, T11), and 83.3 (T2, T8, T11). The average contributions over the ten rounds grouped by MPC are displayed in Figure B2.

It appears that treatments with the same MPC yield quite similar contributions in three of the four comparisons. Note that the treatments within these three comparisons are quite similar to each other. The group sizes never differ by more than 40 subjects and the MPCRs never differ by more than 0.06. The comparison of T1 and T11, however, with 8 versus 100 participants and an MPCR of 0.25 versus an MPCR of 0.02 yields quite different contributions in all of the 10 rounds.

We also include the MPC in Tobit regressions similar to those presented in the main part of the paper. The results are reported in Table B1 and support the idea that MPC is inversely related to contributions. They also reveal that the full model specification based only on MPCR distance yields the best fit based on the AIC values when considering round 1 or round 10 (column (4)). When considering average contributions of all rounds the models including only MPCR distance (columns (3) and (4)) fare better than the models including only MPC (columns (1) and (2)). A model including linear terms and quadratic terms of MPC *and* MPCR distance yields a slightly better fit. In this specification (column (6)) the terms based on MPC are, however, only weakly significant while the terms based on MPCR distance are significant with  $p < 0.05$ .

---

<sup>17</sup> For simplicity we ignore the condition that the number of coalition members must be an integer. This improves the fit of most of the following regression models that include MPC.

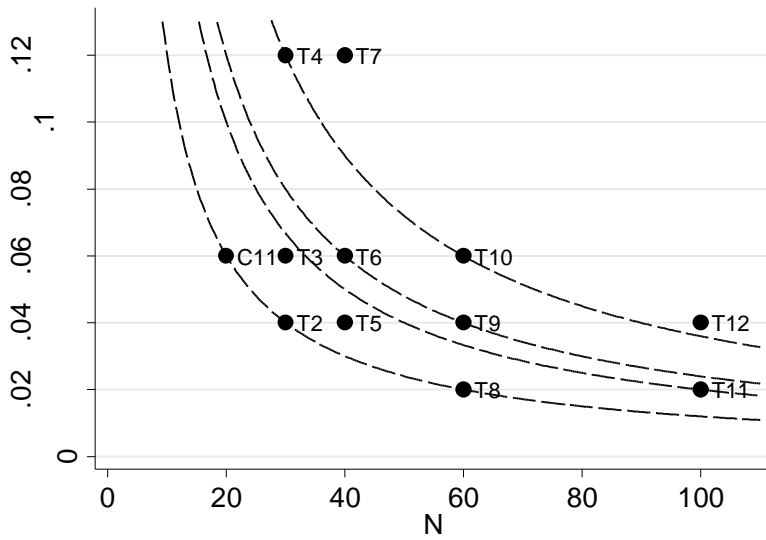


Figure B1: MPCR, group size, and iso-MPC curves (T2 to T12 and C11)

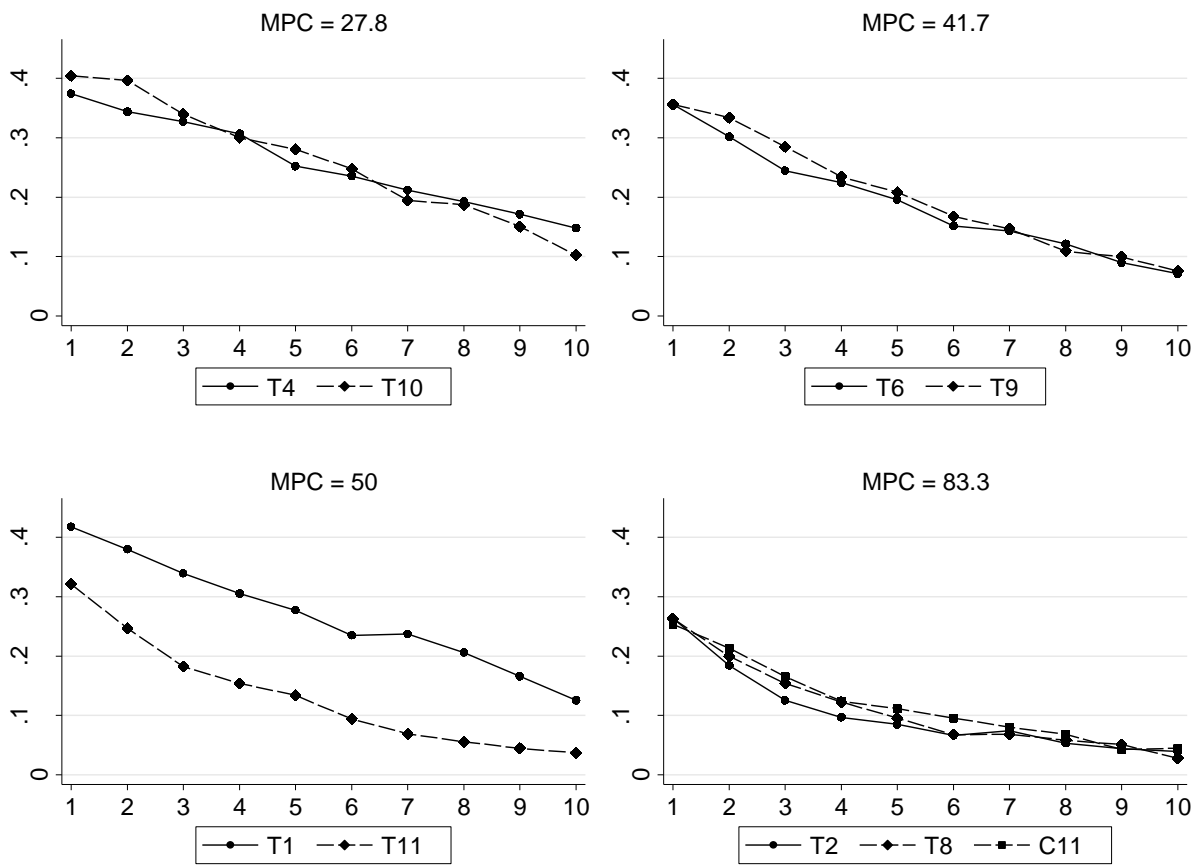


Figure B2: Mean contributions grouped by MPC



	(1)	(2)	(3)	(4)	(5)	(6)
<i>MPC</i>	-0.002*** (0.000)	-0.002 (0.002)			-0.002*** (0.000)	0.001 (0.002)
<i>(MPC)<sup>2</sup></i>		0.000 (0.000)				-0.000 (0.000)
<i>MPCR distance</i>			1.247*** (0.203)	4.555*** (0.649)	0.484* (0.249)	3.503*** (1.285)
<i>(MPCR distance)<sup>2</sup></i>				-31.626*** (5.992)		-23.594** (9.825)
<i>N</i>	4,744	4,744	4,744	4,744	4,744	4,744
<i>N left-censored</i>	839	839	839	839	839	839
<i>N right-censored</i>	443	443	443	443	443	443
<i>AIC</i>	6394.140	6396.083	6421.643	6384.310	6390.718	6384.932

Table B1a – Tobit regressions on contributions in round 1

	(1)	(2)	(3)	(4)	(5)	(6)
<i>MPC</i>	-0.002*** (0.000)	-0.005*** (0.001)			-0.002*** (0.000)	-0.002** (0.001)
<i>(MPC)<sup>2</sup></i>		0.000*** (0.000)				0.000* (0.000)
<i>MPCR distance</i>			1.632*** (0.157)	4.547*** (0.418)	0.938*** (0.184)	3.084*** (0.736)
<i>(MPCR distance)<sup>2</sup></i>				-27.838*** (4.055)		-17.759*** (5.980)
<i>N</i>	4,744	4,744	4,744	4,744	4,744	4,744
<i>N left-censored</i>	540	540	540	540	540	540
<i>N right-censored</i>	15	15	15	15	15	15
<i>AIC</i>	-1648.270	-1675.880	-1617.257	-1751.084	-1737.209	-1764.131

Table B1b – Tobit regressions on average contributions over all rounds

	(1)	(2)	(3)	(4)	(5)	(6)
<i>MPC</i>	-0.001*** (0.000)	-0.004*** (0.001)			-0.000*** (0.000)	-0.001* (0.001)
<i>(MPC)<sup>2</sup></i>		0.000*** (0.000)				0.000* (0.000)
<i>MPCR distance</i>			1.005*** (0.097)	2.027*** (0.277)	0.813*** (0.117)	1.874*** (0.448)
<i>(MPCR distance)<sup>2</sup></i>				-9.578*** (2.611)		-8.775** (3.718)
<i>N</i>	4,744	4,744	4,744	4,744	4,744	4,744
<i>N left-censored</i>	2,689	2,689	2,689	2,689	2,689	2,689
<i>N right-censored</i>	80	80	80	80	80	80
<i>AIC</i>	4294.998	4278.508	4227.447	4211.691	4218.133	4212.625

This table reports average marginal effects on the observed contributions. Regressions include controls for age, gender and location. Standard errors clustered on the session level are given in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-sided).

Table B1c – Tobit regressions on contributions in round 10

## Appendix C

	Coefficient	Standard Error
Treatment T1	0.062	(0.041)
Treatment T2	-0.074**	(0.036)
Treatment T3	0.055	(0.036)
Treatment T4	0.048	(0.036)
Treatment T5	0.000	(0.035)
Treatment T6	0.026	(0.035)
Treatment T7	0.079**	(0.035)
Treatment T8	-0.060*	(0.035)
Treatment T9	0.036	(0.035)
Treatment T10	0.081**	(0.035)
Treatment T11	-0.021	(0.035)
Treatment C3	0.069**	(0.035)
Round <i>t</i>	-0.015***	(0.003)
Treatment T1 * Round <i>t</i>	-0.011***	(0.004)
Treatment T2 * Round <i>t</i>	-0.010***	(0.003)
Treatment T3 * Round <i>t</i>	-0.016***	(0.003)
Treatment T4 * Round <i>t</i>	-0.005	(0.003)
Treatment T5 * Round <i>t</i>	-0.014***	(0.003)
Treatment T6 * Round <i>t</i>	-0.012***	(0.003)
Treatment T7 * Round <i>t</i>	-0.006*	(0.003)
Treatment T8 * Round <i>t</i>	-0.010***	(0.003)
Treatment T9 * Round <i>t</i>	-0.013***	(0.003)
Treatment T10 * Round <i>t</i>	-0.012***	(0.003)
Treatment T11 * Round <i>t</i>	-0.016***	(0.003)
Treatment C3 * Round <i>t</i>	-0.016***	(0.003)
<i>N</i>	47,740	
<i>N</i> left-censored	16,296	
<i>N</i> right-censored	1,572	
AIC	28675.590	

This table reports average marginal effects on the observed contributions per round.

Regressions include controls for age, gender and location. Baseline: Treatment C11. Standard errors are given in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$  (two-sided).

Table C1 – Random-effects Tobit regression of marginal effects on the observed contributions per round